

University of Groningen

Auteursidentificatie

Feringa, Jan

Published in:
Handboek Informatiewetenschap voor bibliotheek en archief

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Feringa, J. (2017). Auteursidentificatie: dat mag geen naam hebben. In G. M. van Trier (editor), *Handboek Informatiewetenschap voor bibliotheek en archief* (blz. 1-12). [IV F 210] Vakmedianet.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Auteursidentificatie: dat mag geen naam hebben

JAN FERINGA

1 Persoonsverwisseling en naamsvarianten

Of het nu om het uitzoeken van een leuk boek in de bibliotheek gaat of het contact leggen met een auteur van een artikel, het is altijd heel vervelend als de juiste auteur niet gevonden kan worden of verwisseld wordt met een andere auteur. Een auteur identificeren we over het algemeen op basis van zijn of haar naam. Nu zal er niet snel verwarring ontstaan bij A.F.Th. van der Heijden, maar bijvoorbeeld bij een naam als J. de Jong bestaat er een aanzienlijke kans op persoonsverwisseling. De risico's daarop zijn nog groter bij auteurs met familienamen met tientallen miljoenen dragers als Wang, Zhang en Li.

Een ander probleem bij het identificeren van auteurs is het voorkomen van meer dan één naamsvariant voor dezelfde persoon. Het bestaan van die varianten kan te maken hebben met het hanteren van één of meer pseudoniemen of door verandering van de naam. Zo is het in een aantal landen gebruikelijk de achternaam aan te passen bij huwelijk en scheiding. Ook ontstaan er varianten door transliteratie vanuit een ander alfabet of vanuit karakters. Denk bijvoorbeeld aan de vele schrijfwijzen van de namen van de voormalige Russische partijleider Хрущёв (Chroesjtsjov, Khrushchev, etc.) of zijn Chinese collega 毛泽东 (Mao Zedong, Mao Tse-toeng).

We zien ook regelmatig dat onbedoeld nieuwe naamsvarianten ontstaan in bestanden of zoekmachines, door het omdraaien van voornaam en achternaam. Dit gaat vooral mis bij achternamen die ook als voornaam voorkomen, zoals Céline, Christie, Cornelis, Gordon, Jens, Martin, Otto, Tiffany, Thomas, etc. Dat in veel Oost-Aziatische talen en ook het Hongaars eerst de familienaam wordt vermeld en daarachter de voornaam, draagt ook bij aan de omdraaiingen. De componist Béla Bartók heet in het Hongaars Bartók Béla.

Ambigüiteit over wie er schuilgaat achter de naam op een boek, artikel, partituur, schilderij, etc. zorgt voor problemen op allerlei vlakken. Deze problematiek speelt bijvoorbeeld bij het maken van titelbeschrijvingen, bij het benade-

ren van experts,¹ bij het zoeken van literatuur, bij het toeschrijven van een kunstwerk, bij het maken van een bibliografie of citatieanalyse, bij het beoordelen van aanvragen voor onderzoekfinanciering, etc.

2 De oplossing: unieke, onveranderlijke codes

Deze ambiguïteit doet zich natuurlijk niet alleen voor bij het identificeren van auteurs. De kans op verwarring tussen twee of meer entiteiten komt op allerlei terreinen voor, bijvoorbeeld bij het uit elkaar houden van consumentenproducten of van burgers van een land. De voor de hand liggende oplossing is dan het opzetten van een organisatie voor het uitdelen en beheren van unieke, onveranderlijke codes. Bij consumentenproducten gebeurt dat door GS1 die wereldwijd GS1 codes (ook wel EAN-nummers genoemd) uitdeelt. Onder dit regime vallen bijvoorbeeld ook de codes voor boeken (ISBN) en tijdschriften (ISSN). Voor personen die een relatie hebben met de Nederlandse overheid, gebeurt iets vergelijkbaars door de Rijksoverheid die burgerservicenummers uitdeelt en beheert.

Ook bij auteurs zouden we graag een unieke, onveranderlijke (persistente) code willen hebben per auteur, die in de hele keten (auteur, uitgever, zoekmachine, lezer, citeerder) gehanteerd wordt en die liefst ook nog in de vorm van Linked Open Data toegankelijk is.² Hiertoe zijn diverse initiatieven ondernomen. Deze kunnen we onderscheiden in nationale initiatieven (zoals in Nederland NTA en DAI), internationale initiatieven uit de commerciële wereld (zoals Researcher-ID, Scopus Author ID en Google Scholar Citation Profile) en internationale initiatieven van organisaties zonder winstoogmerk (zoals ISNI en ORCID).

Naast de geografische dekking en commerciële of niet-commerciële oriëntatie van een code, is een ander belangrijk onderscheid of de codes worden aangevraagd door (of namens) de auteur zelf of dat een daartoe bevoegde instelling dat doet. Als voorbeeld voor de verschillende auteurcodes in dit artikel wordt de recente Nobelprijswinnaar Ben Feringa gebruikt.

3 Auteurcodes in Nederland: van PPN naar DAI.

In Nederland kennen we sinds 1980 de Nederlandse Thesaurus van Auteursnamen (NTA) als onderdeel van het GGC (Gemeenschappelijk Geautomatiseerd Catalogiseersysteem). De NTA is gebaseerd op de persoonsnamen-

1 Cals, Jochen WL and Kotz, Daniel. (2008). Researcher identification: the right needle in the haystack. *The Lancet*, 371(9631), 2152-2153.

2 <https://www.semantic-web.at/LOD-TheEssentials.pdf>

saurus van de bibliotheekcoöperatie OCLC die door Nederlandse bibliotheken wordt aangevuld en bijgehouden. Alle auteurs die in de NTA worden opgenomen, krijgen automatisch een uniek nummer toegekend, het Pica Productie Nummer (PPN). Het PPN telt negen of tien tekens, allemaal numeriek. Het laatste teken is een controleteken, een zogeheten *modulus 11 check digit*¹, net als in het ISBN. Het PPN van Ben Feringa is 067545432.

De NTA werkt met aparte PPN's voor pseudoniemen. Als de auteur ook onder zijn persoonsnaam publiceert, heeft hij of zij dus meer dan één PPN. In de verschillende PPN's van één auteur wordt wel naar elkaar verwezen in kenmerkcodes (kmc) 400 en 410. Zie het voorbeeld hieronder, waarin 068830572 het PPN is voor J.H.F. Grönloh en 097158097 het PPN van zijn pseudoniem Nescio.

PPN 097158097

005 Tpv

100 @Nescio

400 pseud. van: !068830572!Jan Hendrik Frederik Grönloh

900 1882-1961

900 letterkundige

PPN 068830572

005 Tpv

006 /1D

007 /1nl

100 J.H.F.@Grönloh

110 Jan Hendrik Frederik@Grönloh

200 Frits@Grönloh

300 1882-1961

410 zie ook: !097158097!Nescio

710 letterkundige

710 zakenman

790 =A n 83200403

791 =A 0000000120278745 =2 isni

900 Med. NLMD

900 NLMD/G 877

Andere versies of weergaves van de gehanteerde auteursnaam, incl. meisjesnamen, worden in het auteursrecord zelf verwerkt d.m.v. een variant, zie ook kmc 110 en 200 hierboven. Zo telt het PPN van de componist Pjotr Iljitsj Tsjaikovski meer dan dertig naamsvarianten. Zie hieronder zijn GGC-record

1 <http://www.pgrocer.net/Cis51/mod11.html>

met alle varianten in kmc 200. Het samenbrengen van al deze varianten onder één PPN heeft als voordeel dat alle werken van deze componist via één zoekactie teruggevonden kunnen worden. Het is in het GGC overigens niet de bedoeling dat aparte auteursrecords gehanteerd worden voor de verschillende rollen van een auteur (schrijver, componist, schilder, etc.).

PPN 068877110
 005 Tpx
 007 /1su
 110 [\01,ba\]Petr Il'ič@Čajkovskij
 110 [\01,cr\]Петр Ильич@Чайковский
 200 Peter I.@Tschaikowsky
 200 P.I.@Čajkovskij
 200 Piotr Ilyich@Tchaikovsky
 200 Peter Iljitsch@Tschaikowsky
 200 Peter Ilyich@Tchaikovsky
 200 Piotr Ilyitch@Tchaikovski
 200 Pjotr Iljitsj@Tsjaikofski
 200 Piotr@Ciaikovsky
 200 Pjotr Iljitsj@Tsjaikovski
 200 Pyotr Ilyich@Tchaikovsky
 200 P.@Tschaikowski
 200 P.@Tschaikowsky
 200 Peter Iljitsj@Tsjaikowski
 200 Pjotr Iljitsj@Tsjaikowskij
 200 Pyotr Il'yich@Tchaikovsky
 200 Peter@Tschaikowsky
 200 Peter Illitsj@Tsjaikowsky
 200 P.I.@Tschaikowsky
 200 Pjotr Il'yich@Tchaikovsky
 200 Piotr Ilic@Ciaikovskij
 200 Peter Ilyitch@Tschaikovsky
 200 Pyotr Il'yich@Tchaikovsky
 200 @Tsjaikovski
 200 @Tsjaikovsky
 200 P.I.@Tsjaikovski
 200 @Csajkovszkij
 200 @Chaikovsky
 200 Peter J.@Tschaikowsky
 200 Peter@Tchaikovsky
 200 Piotr@Czajkowski
 300 [\01,ba\]1840-1893
 300 [\01,cr\]1840-1893

791 =A 0000000121449518 =2 isni
 900 componist
 900 mb
 900 +
 910 Tsjaikovski

Helaas ontbraken in de NTA veel wetenschappelijke auteurs, omdat auteurs van artikelen in wetenschappelijke tijdschriften en hoofdstukken in bundels vaak niet in het GGC gecatalogiseerd en dus ook niet gethesaureerd werden. Vanaf 2005 is daarom een additionele stroom van auteurs richting de NTA op gang gekomen vanuit de Nederlandse wetenschappelijke instellingen. Deze instellingen registreren namelijk allemaal hun research output in zogenaamde Current Research Information Systems (CRIS).¹ Vanuit Metis, een CRIS dat in die tijd door alle universiteiten gebruikt werd, kregen wetenschappers een zogenaamde DAI (Digital Author Identifier) toegekend. Als een wetenschapper al een PPN had, dan 'promoveerde' het PPN tot DAI in het CRIS en andere systemen die gebruik maken van de gegevens uit het CRIS (institutioneel repository, NARCIS,² Google Scholar, etc.). Als een wetenschapper nog geen PPN had, werd de DAI vanuit Metis geïmporteerd in de NTA.

Begin 2017 bevat de NTA ruim 2,6 miljoen personen, waarvan ruim 80.000 DAI's. De PPN's in de NTA worden sinds enkele jaren gesynchroniseerd met de internationale ISNI auteurcode (zie paragraaf 5.1), in eendrachtige samenwerking door de Koninklijke Bibliotheek (KB) en OCLC. Zie ook kmc 791 in het voorbeeld hierboven. Eerst gebeurde de synchronisatie indirect via VIAF, een internationale online dienst die thesaurusbestanden koppelt.³ Sinds kort is de NTA echter een zelfstandige bron voor ISNI. Voor bibliotheken die niet meer in het GGC catalogiseren maar rechtstreeks in WorldCat, bestaan aparte afspraken met de KB over het proces van het toekennen van een ISNI aan een auteur.

Sterk punt van de NTA is de hoge kwaliteit van het proces van zoeken en – indien noodzakelijk – toekennen van een nieuwe PPN aan een auteur, doordat dit alleen door goed geschoold personeel gebeurt. Hiermee onderscheidt de NTA zich van systemen die gebaseerd zijn op algoritmes (Scopus, Google Scholar) en/of desambiguering door de auteur zelf (Scopus, Web of Science, Google Scholar). Dat het proces van toekennen van een PPN onafhankelijk is

1 Feringa, J. (2014). CRIS – Current Research Information Systems. In *Handboek Informatiewetenschap voor bibliotheek en archief*. (p. 1). [IV A 115] Alphen aan de Rijn: Vakmedianet.

2 <http://www.narcis.nl/about/Language/nl>

3 <http://www.informatieprofessional.nl/bijdragen/2013/03/backoffice-op-de-voorgroen/index.xml>

van de inbreng van de auteur zelf, komt bijvoorbeeld goed van pas als een DAI of PPN postuum moet worden toegekend.

Toch werd, ondanks de toevoeging van veel wetenschappelijke auteurs, steeds duidelijker dat aan de NTA te veel bezwaren kleefden. Zwakke punten van de NTA zijn vooral de gerichtheid op Nederland en de nadruk op auteurs van schriftelijke werken. Mede vanwege de toenemende internationalisering van de wetenschap en de groeiende behoefte aan het geautomatiseerd identificeren van de correcte auteurs van een werk, gingen deze beperkingen steeds zwaarder wegen.

4 De aanpak van aanbieders van commerciële academische zoekmachines

4.1 *Scopus*

Scopus is een commerciële bibliografische database met samenvattingen en citaties van tientallen miljoenen wetenschappelijke artikelen en ruim honderd-duizend boeken. Elsevier bracht Scopus in 2004 op de markt en sinds 2006 werkt Scopus met auteurcodes, de Scopus Author ID, bestaande uit 10 of 11 cijfers. Scopus bepaalt zelf op basis van algoritmes deze ID. Een auteur loopt dus, zeker als er meerdere naamsvarianten in omloop zijn, het risico meerdere Scopus Author ID's te hebben en ook kunnen er publicaties aan de verkeerde ID gekoppeld zijn. Dit kan een auteur handmatig herstellen, ook in combinatie met de ORCID auteurcode (zie ook paragraaf 5.2).^{1,2} Het is dus niet mogelijk een Author ID bij Scopus aan te vragen. Scopus bevat inmiddels 12,9 miljoen Author ID's met 2 of meer publicaties. De Scopus Author ID van Ben Feringa is 35513908800.³

Omdat Scopus zich hoofdzakelijk richt op auteurs van wetenschappelijke artikelen en bovendien niet onafhankelijk is, is de toepasbaarheid van de Scopus Author ID buiten Scopus beperkt.

4.2 *Web of Science*

Web of Science (WoS) is een commerciële onlinedienst voor wetenschappelijke citatie-indexering, tot voor kort van Thomson Reuters en inmiddels van Clarivate Analytics. WoS geeft toegang tot meerdere databases, waarmee ge-

1 http://help.elsevier.com/app/answers/detail/a_id/2321/p/8150/related/1

2 <https://www.elsevier.com/solutions/scopus/support/authorprofile>

3 <https://www.scopus.com/freelookup/form/author.uri>

zocht kan worden in allerlei gespecialiseerde subdomeinen van wetenschappelijke vakgebieden.

WoS kent vanaf 2008 een eigen identificatiesysteem voor auteurs, de ResearcherID. Waar Scopus zelf de auteurcode genereert en koppelt aan publicaties, laat WoS het initiatief aan de auteur. Deze moet eerst een ResearcherID aanvragen¹ en kan daarna publicaties ‘claimen’.

De structuur van een ResearcherID is een letter, gevolgd door een koppelteken, vier cijfers, opnieuw een koppelteken en vervolgens het jaar van aanvraag. WoS bevat op dit moment ca. 800.000 ResearcherID’s. De ResearcherID van Ben Feringa is C-1665-2013.² Net als Scopus ondersteunt WoS de ORCID auteurcode en is de toepasbaarheid van de ResearcherID buiten WoS beperkt.

4.3 Google Scholar

Google Scholar is een zoekmachine speciaal voor wetenschappelijke literatuur in 2004 door Google op de markt is gebracht. Google Scholar is – in tegenstelling tot Scopus en WoS – gratis, staat bekend om de goede dekking van Engelstalige wetenschappelijke tijdschriften, maar ontsluit bijvoorbeeld ook boeken, grijze literatuur, rechterlijke uitspraken, patenten en software.³ Waar WoS en Scopus alleen publicaties meetellen die aan bepaalde criteria voldoen, legt Google Scholar de lat veel lager. Google Scholar levert daarom vaak meer publicaties op bij dezelfde zoekvraag dan Scopus en WoS en vaak ook met hogere aantallen citaties, hoewel verwisseling door Google Scholar van velden, zoals voor- en achternaam, soms ook tot een lager aantal citaties leidt.⁴

Sinds 2012 kent Google Scholar zogenaamde Citation Profiles. Hierin kan een auteur zijn of haar publicaties plus de citaties van deze publicaties bijhouden, alsmede bibliometrische ratio’s zoals h-index en i10-index.⁵ Hoewel veel gebruikt, zijn deze Citation Profiles alleen bruikbaar binnen het Google Scholar

1 <http://www.researcherid.com/SelfRegistration.action>

2 <http://www.researcherid.com/ViewProfileSearch.action>

3 Fagan, Jody Condit (2017). An Evidence-Based Review of Academic Web Search Engines, 2014-2016: Implications for Librarians’ Practice and Research Agenda. *Information Technology and Libraries*, 36(2) 27.

4 <http://www.harzing.com/pophelp/accuracy.htm>

5 De Hirsch-index of h-index is een indicator voor de ‘carrière-impact’ van een wetenschappelijk onderzoeker op basis van zijn of haar publicaties en bijbehorende citaties. Aan het vergelijken van onderzoekers op basis van hun h-index kleven risico’s. Zowel de keuze van de bron van publicaties en citaties (bv. WoS of Google Scholar), het vakgebied en de leeftijd van de auteur zijn van grote invloed op de hoogte van de h-index. De i10-index is een andere indicator voor de prestaties van een onderzoeker. Deze indicator geeft het aantal publicaties weer met minstens 10 citaties bij Google Scholar.

ecosysteem en vaak alleen zichtbaar voor de auteur zelf. Ze voorzien dus niet in de behoefte van een unieke, onveranderlijke auteurcode.

Naast de drie belangrijke bronnen van wetenschappelijke literatuur hierboven, zijn er nog meer databases die met auteurcodes werken. Het gaat vaak om domeinspecifieke databases zoals MathSciNet (wiskunde).

5 Onafhankelijke globale auteurcodes

5.1 ISNI

ISNI staat voor International Standard Name Identifier en is een standaard ontwikkeld onder auspiciën van ISO (International Organization for Standardization) voor de unieke identificatie van partijen die betrokken zijn bij de creatie, productie, beheer en distributieketens van (kunst)werken.¹ ISNI positioneert zich als ‘bridge identifier’ en richt zich voornamelijk op het integreren van bestaande databases en *authority files* zoals de NTA. Deze standaard staat dus dicht bij de bibliotheekwereld en zijn kwalitatief hoogwaardige bronbestanden. De centrale kwaliteitscontrole van ISNI is belegd bij de Bibliothèque National de France en de British Library.²

De ISNI-standaard is vastgesteld in 2012 en bevat al 8,7 miljoen auteurs en ca. 600.000 organisaties. De ISNI telt zestien tekens, meestal weergegeven in vier blokken van vier. De eerste vijftien tekens zijn numeriek. Het laatste teken is een controleteken, conform ISO/IEC 7064:2003, MOD 11-2 en kan numeriek zijn of een ‘X’.³

Een ISNI kan in de Verenigde Staten tegen een vergoeding van \$ 25 door een individueel persoon worden aangevraagd,⁴ maar gebruikelijk is aanmelding via een zogenaamde ‘Registration Agency’, in Nederland de KB.⁵ Opvallend is dat het toegestaan is dat één persoon of partij meerdere ISNI’s gebruikt, bijvoorbeeld een aparte ISNI voor een pseudoniem (naast of in plaats van een ISNI met de officiële naam van de auteur) of één ISNI als schrijver en een andere ISNI als musicus. ISNI’s behorend bij dezelfde persoon of partij kunnen gelinkt worden. De ISNI van Ben Feringa is 0000 0001 1653 2165.

De samenwerking tussen brancheorganisaties, zoals BUMA/STEMRA, en de KB over ISNI moet nog op gang komen. Er zijn wel al veel ISNI’s van Neder-

1 http://www.iso.org/iso/catalogue_detail?csnumber=44292

2 https://www.surfspace.nl/media/pdfs/260c2f907c2bd0eff492eeaf4fcb244f_2015-05-19-nota-toekomst-nederlandse-infrastructuur-voor-auteursidentifiers.pdf

3 <https://www.iso.org/obp/ui/#iso:std:iso-iec:7064:ed-1:v1:en>

4 <http://numericalgurus.com/isni-individual-registration/>

5 <http://www.isni.org/do-you-have-an-isni>

landse musici ingelezen vanuit internationale bronnen als <http://musicbrainz.org/> en <http://www.scapr.org/ipd> zoals de ISNI van Mariska Veres:

ISNI: 0000 0001 1765 9344

Name: **Mariska Veres** (Dutch singer of Shocking Blue)
Mariska Veres (holenderska wokalistka rockowa)
Mariska Veres (Nederlandse zangeres van Shocking Blue)
Mariska Veres (niederländische Sängerin)
Veres, Mariska
Μαρίσκα Βέρες
Маришка Верес
 מריסקה ורס
 ماریسکا فیرس (خواننده هلندی)
 마리스카 베레스
 マリスカ・ヴェレス

Dates: 1949-2006

Creation class: Musical sound recording

Creation role: performer

Related names: Beek, van der Cor (co-performer)
 Leeuwen, van Robbie (co-performer)
 Shocking Blue (isMemberOf)
 Shocking blue (see also from)
 Wal, van der Klaasje (co-performer)

5.2 ORCID

ORCID (Open Researcher and Contributor ID) is een auteurcode die wordt beheerd door de non-profit organisatie ORCID, Inc.¹ De structuur van ORCID is identiek aan de structuur van ISNI. ISNI heeft voor ORCID een blok codes gereserveerd van 0000 0001 5000 0007 tot 0000 0003 5000 000. Uit deze range worden sinds eind 2012 ORCID codes uitgedeeld aan wetenschappers. ORCID kan dus ook beschouwd worden als een subset van ISNI.^{2,3} Ben Feringa heeft, net als menig ander wetenschapper, naast een ISNI ook een ORCID: 0000 0003 0588 8435.

1 <https://orcid.org/>

2 <http://about.orcid.org/content/what-relationship-between-isni-and-orcid>

3 https://web.archive.org/web/20130304210352/http://www.isni.org/isni_and_orcid

ORCID is specifiek gericht op het koppelen van onderzoekers aan hun ‘scholarly research output’. Het gaat dus niet alleen om publicaties, maar bijvoorbeeld ook om data, software, patenten en grantproposals.

Het aanvragen van een ORCID is gratis en simpel.¹ De onderzoeker moet dit echter zelf doen en daarom kan de ORCID dus niet postuum toegewezen worden.² De organisatie nodigt onderzoekers uit om ook hun publicaties te laden in ORCID. Dit kan op drie manieren: via elektronische databases zoals CrossRef, Scopus en WoS, via het uploaden van een bestand met metadata (BibTeX) of met de hand. De meningen verschillen of het verstandig is om publicaties te laden bij ORCID. Enerzijds kan de zichtbaarheid van de publicaties in ORCID de onderzoeker meer bekendheid geven, anderzijds ontstaat er wel weer een extra ‘publicatielijst’ die onderhouden moet worden naast het institutionele CRIS-systeem en naast online netwerken waarop wetenschappers graag actief zijn, zoals ResearchGate³ en Academia.edu⁴.

Er zijn op dit moment ruim 2,9 miljoen ORCID’s aangemaakt en het aantal neemt snel toe. Deze snelle groei wordt bevorderd door de omstandigheid dat belangrijke partijen in het domein van de wetenschap zich achter ORCID geschaard hebben. Steeds meer wetenschappelijke uitgevers, zoals PLoS, Wiley en IEEE, willen de ORCID van de auteur(s) weten bij het indienen van een publicatie.⁵ Ook steeds meer onderzoeksfinanciers, zoals de Wellcome Trust, Autism Speaks (VS), Research Councils UK en het Swedish Research Council, accepteren geen aanvragen meer zonder de ORCID’s van de aanvragers. Tevens zetten belangrijke elektronische databases als WoS, Scopus en CrossRef sterk in op integratie van ORCID’s in hun systemen.

De sleutel van de toekomst van ORCID ligt vooral in handen van de uitgevers. Als alle belangrijke uitgevers aanlevering van de ORCID bij het indienen van een artikel verplicht stellen, dan zullen deze codes kort daarna ook als metadata verschijnen in de elektronische databases. Als dat eenmaal goed werkt, dan wordt het pas mogelijk CRIS-systemen automatisch te vullen vanuit deze elektronische databases, zonder dat auteurs door elkaar gehaald kunnen worden.⁶ Vanuit die CRIS-systemen kunnen dan websites van de instellingen en persoonlijke websites automatisch gevuld worden en kunnen bijvoorbeeld NARCIS, Google Scholar en WorldCat publicatiemetadaten harvesten met

1 Bokdam, G.J. (2016). Auteursidentificatie voor onderzoekers met ORCID. *Pictogram*, 18(1), 6-7.

2 <https://orcid.org/register>

3 <https://www.researchgate.net/>

4 <https://www.academia.edu/>

5 <http://orcid.org/content/requiring-orcid-publication-workflows-open-letter>

6 Dit is extra belangrijk in onderzoeksgebieden waar het gebruikelijk is honderden of duizenden auteurs op te voeren, zie bijvoorbeeld de 5154 auteurs in dit artikel in *Physical Review Letters*: <http://inspirehep.net/record/1356276>.

correcte auteursgegevens. Vooral de mogelijkheid om WorldCat te laten harvesten is erg interessant omdat WorldCat een belangrijke bron is van Linked Open Data.¹

6 Aanbevelingen

Auteurcodes zijn handige hulpmiddelen bij het uit elkaar houden van auteurs. Hiervoor zijn verschillende systemen bedacht. Met name wetenschappelijke auteurs hebben hier veel keus in.² Iemand als Ben Feringa bijvoorbeeld, beschikt over een heel assortiment aan auteurcodes:

PPN	067545432
Scopus Author ID	35513908800
ResearcherID	C-1665-2013
ISNI	0000 0001 1653 2165
ORCID	0000 0003 0588 8435
VIAF	59976288
LCCN ³	nb99166949
GND ⁴	172085551
SUDOC ⁵	067270492

Deze lijst geeft wel aan dat we nog niet zo ver zijn dat we een auteur kunnen identificeren met één unieke, onveranderlijke code. Uit voorgaande paragrafen blijkt dat om verschillende redenen de vele nationale en commerciële standaarden deze sleutelrol uiteindelijk niet zullen kunnen spelen. In de wetenschappelijke wereld krijgt de ORCID, een internationale, niet-commerciële identifier, steeds meer momentum als dé auteurcode. Voor andere auteurs als schrijvers, musici, componisten, kunstenaars, etc. is het duidelijk dat ISNI (ook internationaal en niet commercieel) de standaard van de toekomst is. Beide systemen kennen echter nog haken en ogen en ook de organisatie rond beide standaarden kent nog gebreken. Om de transitie naar en gebruik van

1 <https://www.oclc.org/en/WorldCat/data-strategy.html>

2 Op <http://libguides.utsa.edu/AuthorProfile/startstaat> een overzichtelijke tabel waarin ORCID, Researcher ID, Scopus Author ID en Google Scholar Citations op diverse aspecten met elkaar worden vergeleken.

3 Library of Congress Control Number; zie ook kmc 790 bij Nescio in hoofdstuk 3.

4 Gemeinsame Normdatei; standaard die in Duitstalige landen veel gebruikt wordt.

5 Système universitaire de documentation; gezamenlijke catalogus voor het Franse hogere onderwijs.

deze standaarden te vergemakkelijken, staat hieronder een aantal aanbevelingen.

Voor universiteiten en andere instellingen voor hoger onderwijs:

- Het is belangrijk dat alle onderzoekers een ORCID hebben en die bij al hun uitingen en research output weergeven. Laat dit uitgangspunt terugkomen in de indienstredingsprocedures, universitaire systemen (zoals CRIS en het personeelssysteem), profielpagina's, reviews, onderzoeksaanvragen, etc.
- Zorg dat ondersteunende diensten, zoals de bibliotheek, funding office, research data office en clinical research office, de wetenschappers actief informeren over de mogelijkheden en voordelen van ORCID en over de stand van zaken betreffende ORCID bij de verschillende uitgevers, funders en andere belangrijke partijen.
- Zorg dat de ORCID's van de wetenschappelijke medewerkers in het lokale CRIS goed ingevuld zijn en dat het CRIS wordt geharvest door WorldCat, Narcis en OpenAIRE (geaggregeerde literatuur repository van de EU).¹

Wetenschappers moeten zo spoedig mogelijk een ORCID aanvragen en die vermelden bij al hun toekomstige en (voor zover dat mogelijk is) bestaande onderzoeksresultaten (data, software, patenten, etc.) en de bijbehorende uitingen. Verder is het verstandig om op de website van ORCID te checken of eventuele vermeldingen als (co-)auteur wel kloppen en zo niet, die vermeldingen te verbeteren.

Voor bibliotheken die in het GGC catalogiseren verandert er voorlopig niets, zij kunnen de NTA blijven gebruiken. Voor bibliotheken die in WorldCat catalogiseren geldt een aparte procedure rond het toekennen van ISNI's in samspraak met de KB. Voor andere organisaties (van musici, beeldend kunstenaars, etc.) is het aan te raden zo snel mogelijk met de KB een goede werkwijze over het vastleggen van ISNI's overeen te komen.

Voor individuele schrijvers, musici, kunstenaars, etc. is het verstandig om op de website van ISNI² te checken of eventuele vermeldingen als (co-)auteur wel kloppen en zo niet, die te verbeteren. Als nog geen ISNI is toegekend, dan kan men contact opnemen met een Registration Agency³. Voor Nederland is dat de KB, te benaderen via Catalogisering@kb.nl.

1 <http://api.openaire.eu/>

2 <http://www.isni.org/search>

3 <http://www.isni.org/do-you-have-an-isni>